

Ripple Effects of Data Sharing: Ethical Considerations

Richard R. Sharp, PhD

Director of Bioethics Research, Cleveland Clinic

Co-Director, Center for Genetic Research Ethics & Law, Case Western

Jalayne J. Arias, JD

Fellow, Cleveland Fellowship in Advanced Bioethics, Cleveland Clinic

Overview

- Highlight potential benefits and harms of expanded data sharing
- Discuss participant perspectives on secondary data uses
- Offer some suggestions for addressing ethical challenges

Motivations for Data Sharing

“...science and creativity are furthered by access to openly available data, and that data created by publicly funded bodies should be freely available in the research community.”

Kaye Jane, Heeney C, Hawkins N, de Vries J, Boddington. Data sharing in genomics-re-shaping scientific practice. Nature Rev 2009; 10:331.



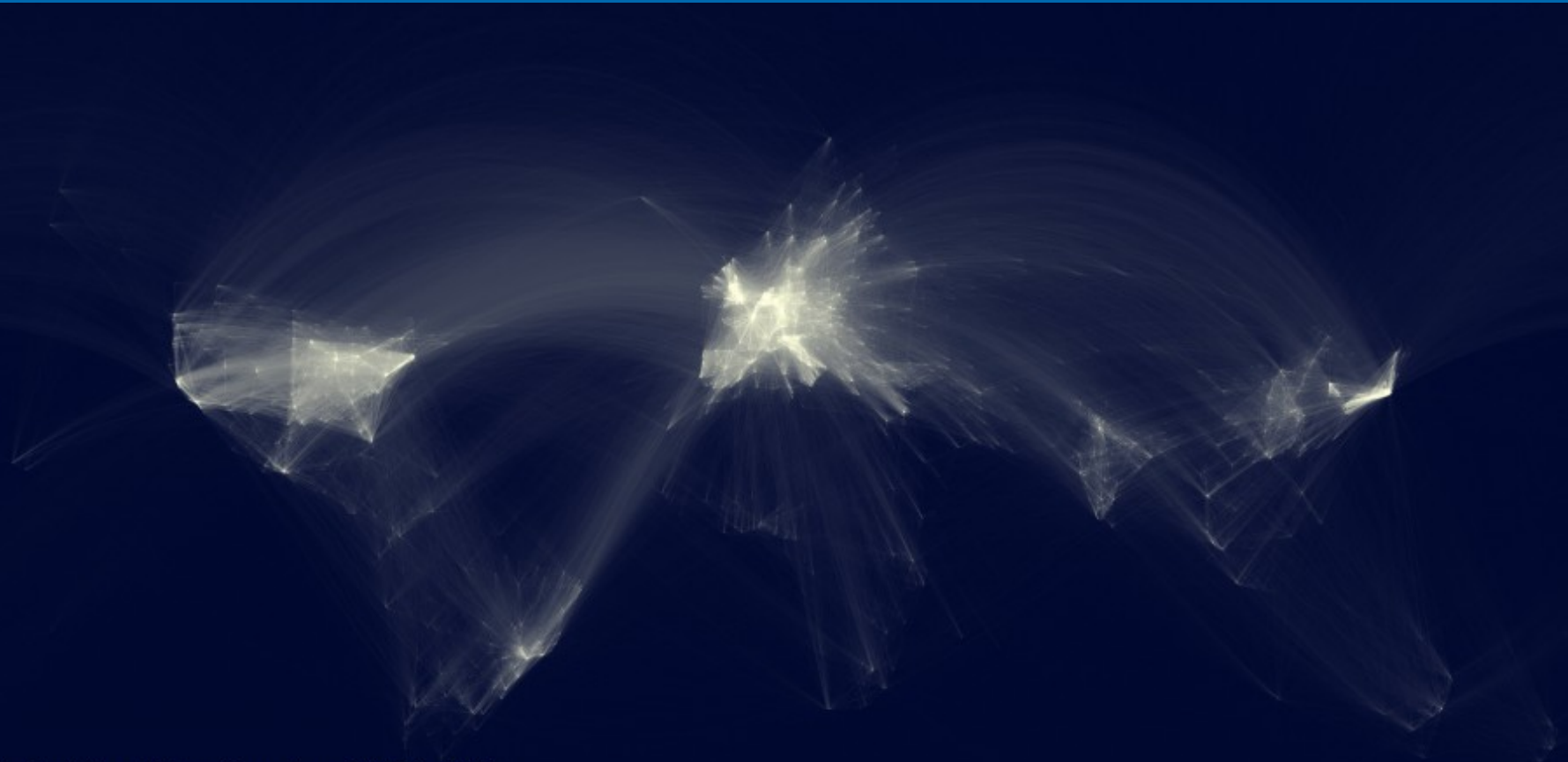
Goals of Data Sharing

Increase value of data for:

- Social and public health through greater access
- Economic yield of federal investment
- New data analysis and secondary confirmation of research findings



The Globalization of Science



Map of scientific collaborations from 2005 to 2009

Computed by Olivier H. Beauchesne @ Science-Metrix, Inc.

Data from Scopus, using books, trade journals and peer-reviewed journals

<http://flowingdata.com/2011/01/27/map-of-scientific-collaboration-between-researchers/>

Scientific Commons: Data Sharing as a New Norm

Sharing of Data Leads to Progress on Alzheimer's

Published: August 12, 2010

New York Times

<http://www.nytimes.com/2010/08/13/health/research/13alzheimer.html>

“THE KEY TO THE ALZHEIMER’S PROJECT WAS AN AGREEMENT AS AMBITIOUS AS ITS GOAL: NOT JUST TO RAISE MONEY, NOT JUST TO DO RESEARCH ON A VAST SCALE, **BUT ALSO TO SHARE ALL THE DATA**, MAKING EVERY SINGLE FINDING PUBLIC IMMEDIATELY, AVAILABLE TO ANYONE WITH A COMPUTER ANYWHERE IN THE WORLD.

NO ONE WOULD OWN THE DATA. NO ONE COULD SUBMIT PATENT APPLICATIONS, THOUGH PRIVATE COMPANIES WOULD ULTIMATELY PROFIT FROM ANY DRUGS OR IMAGING TESTS DEVELOPED AS A RESULT OF THE EFFORT.”



Concerns about Expanded Data Sharing: The Ripple Effects



http://scienceblogs.com/startswithabang/2010/08/what_is_the_speed_of_gravity.php

Participant Perspectives

- Adequacy of prospective informed consent
- Evaluation of non-conventional risks, e.g.
 - Intrinsic or “dignitary” harm to subjects, e.g. persons opposed to abortion might unwittingly participate in research that identifies a gene linked to a non-lethal birth defect—and that finding may increase the number of elective abortions)
 - Consequential harms, e.g. loss of privacy or identification of an individual using other techniques

These considerations suggest a need to extend greater levels of control to participants after data collection.

How Bioethics Research Can Help

- Improve communication with participants
 - Clarify key information to be discussed during the informed consent process
 - Anticipate common questions or concerns about data sharing
- Clarify patient views about the control and disposition of primary data and stored biological materials
- Define potential barriers to research recruitment

Almost no research has been done examining public views of data sharing

Table 2. Consent for Research on Stored Samples*

	Consent Necessary for Clinically Derived Samples?		Consent Necessary for Research-Derived Samples?	
	Personally Identified	Anonymized	Personally Identified	Anonymized
Overall (N = 504)	65.8	27.3	29.0	12.1
Cohort				
In research on Alzheimer disease (n = 246)	68.6	22.8	24.4	8.1
Medicare beneficiaries (n = 258)	63.1	29.8	33.5	15.1
Sex				
Male	60.6	22.3	24.2	7.6
Female	69.0	29.0	32.0	14.3
Age, y				
50-54	83.8†	34.2	21.1	5.3
55-64	69.5	19.8	28.7	12.2
65-74	64.0	28.7	28.2	12.0
≥75	58.9	26.9	33.3	12.6
Income, \$				
<25 000	66.7	36.4	34.9	17.8
25 000-75 000	60.9	24.9	21.8	8.1
>75 000	67.6	20.0	26.7	9.5
Education				
≤High school	67.4	37.0	35.8	17.4
Some college or college graduate	61.7	22.8	25.5	11.0
At least some graduate education	69.4	22.1	27.5	8.0
Race				
White	67.2	26.3	28.7	11.0
African American	47.6‡	18.2	31.8	18.2
Hispanic	50.0	25.0	12.5	0

*Data are percentage of respondents who stated their consent should be required for research using 4 different types of stored samples originally obtained from them.

†Older individuals are significantly less likely to state that their consent should be required.

‡Nonwhites are significantly less likely to state that their consent should be required.

Table 2. Stratification of Authorization of Unlimited Future Research by the Characteristics of the 890 Participants

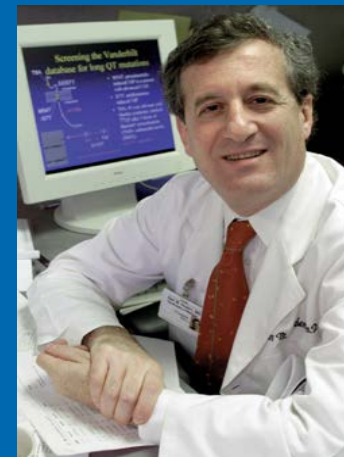
Characteristic	No. of Participants Given the Option	% Choosing the Option (95% CI)*	P Value†
Overall	890	87.1 (84.7-89.2)	NA
Sex			
Male	416	88.9 (85.5-91.8)	.12
Female	474	85.4 (81.9-88.5)	
Age, y			
0-17‡	102	92.2 (85.1-96.6)	.10
≥18	788	86.4 (83.8-88.7)	
Race			
White	722	88.4 (85.8-90.6)	.002
African American	84	75.0 (64.4-83.8)	
Other	84	88.1 (79.2-94.1)	
Residence			
Virginia, Maryland, or District of Columbia	492	85.2 (81.7-88.2)	.06
Other part of the United States	378	88.9 (85.3-91.9)	
International	20	100.0 (83.2-100.0)	
Participant type			
Affected individual	471	86.4 (83.0-89.4)	.66
Family member	179	86.6 (80.7-91.2)	
Healthy volunteer	240	88.8 (84.1-92.5)	
Study type			
Prospect of direct medical benefit	216	88.0 (82.9-92.0)	.66
No prospect of direct medical benefit	674	86.8 (84.6-89.8)	

Abbreviations: CI, confidence interval; NA, data not applicable.

*The exact binomial 95% CI is given.

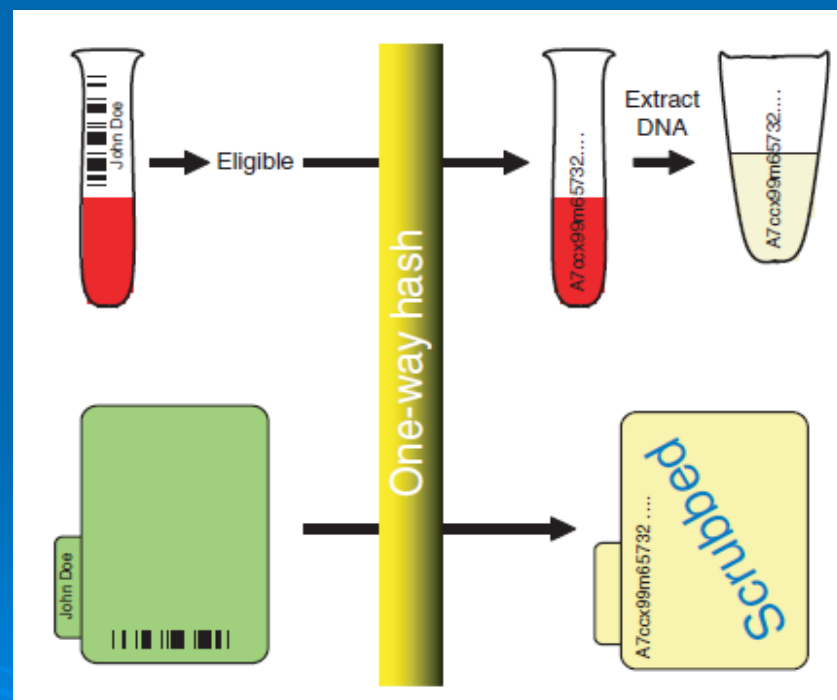
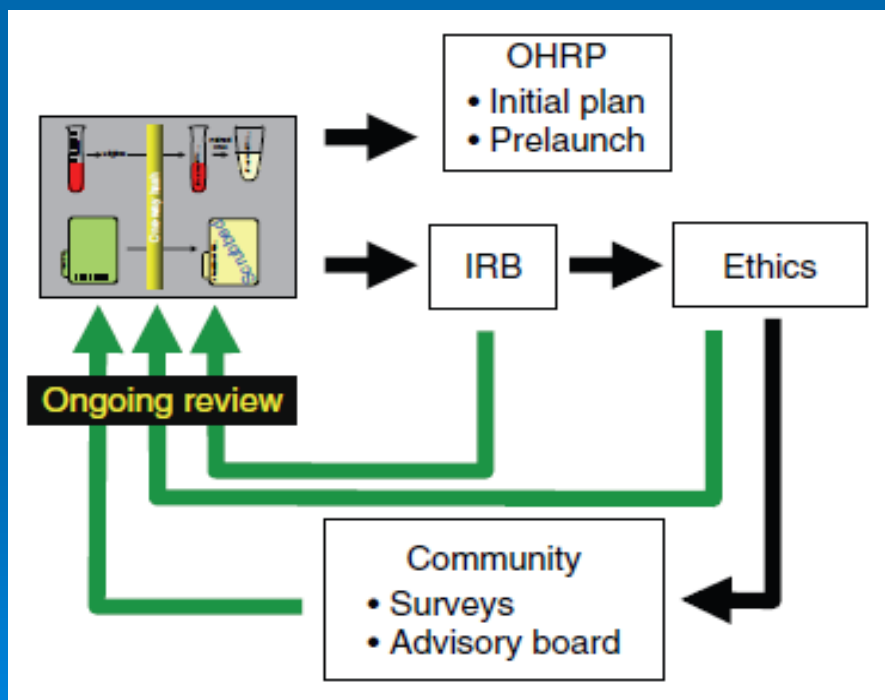
†P values obtained using the χ^2 test for independence. The following P values were obtained after excluding the international participants: sex, .12; age, .10; race, .003; residence, .11; participant type, .56; and study type, .68.

‡For this group (considered children), the decision was made by parents.



Development of a Large-Scale De-Identified DNA Biobank to Enable Personalized Medicine

DM Roden¹⁻³, JM Pulley⁴, MA Basford^{1,4}, GR Bernard^{2,4}, EW Clayton^{5,6}, JR Balser^{3,4} and DR Masys⁷

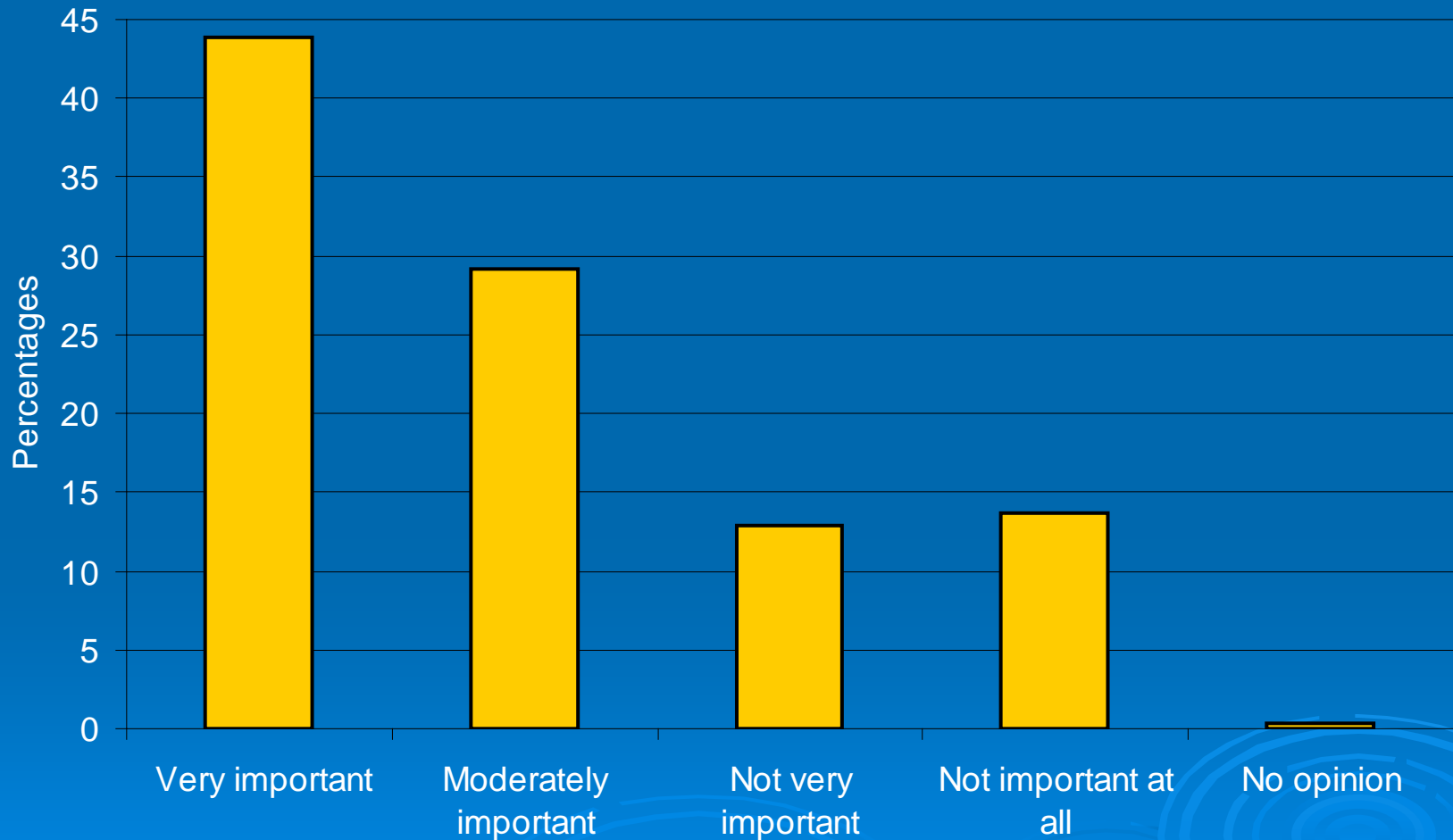


Views on Future Use of Biological Materials

Okay to study different diseases	79%
Willing to sign a one-time release	73%
Okay for different researchers to use	61%

Hull SC, *Am J Bioeth* 2009

Disclosure of Secondary Use of Stored Biological Materials



Preferences For Notification Vs. Permission

	If Anonymous	If Identifiable
Notification Required	43%	42%
Permission Required	56%	57%

Patients do not view distinctions between identifiable, coded, double-coded, and anonymous samples as important to decisions about sample donation.

Potential Privacy Threats

- Individual Identification and Re-Identification
 - These are rare events; people are increasingly aware of numerous threats involving “informational privacy”
- Participants should be informed about privacy protections, but these are often over-emphasized in human subjects protections
- Data may be available to users who are not subject to institutional safeguards, oversight, and professional codes of conduct

Potential Harms: Data Integrity

- Prepublication data may be released prior to quality control or full analysis
- Problematic interpretations by special interest groups
- Controversial or ethically objectionable analysis of existing datasets (data mining of GWAS data for so-called “gay genes”)

Potential Harms: Data Integrity

Table 2

Findings Supported by the Use of the HGP Diversity Panel

Type	Publication	Excerpt from Abstract	Interpretation
Addiction	Bierut, L. J., et al. (2008). "Variants in nicotinic receptors and risk for nicotine dependence." <i>Am J Psychiatry</i> 165(9): 1163-71.	"A genetic variant marking an amino acid change showed association with the smoking phenotype ($p = 0.007$)...t its frequency varied across human populations (0% in African populations to 37% in European populations)."	Europeans are More Susceptible to Nicotine Dependence [29]
Ancestry	Need, A. C., et al. (2009). "A genome-wide genetic signature of Jewish ancestry perfectly separates individuals with and without full Jewish ancestry in a large random sample of European Americans." <i>Genome Biol</i> 10(1): R7.	".. within Americans of European ancestry there is a perfect genetic corollary of Jewish ancestry which, in principle, would permit near perfect genetic inference of Ashkenazi Jewish ancestry."	Jewish People are Genetically Distinct [30]
Genetic Variation	Rosenberg, N. A., et al. (2002). "Genetic structure of human populations." <i>Science</i> 298(5602): 2381-2385.	"...without using prior information about the origins of individuals, we identified six main genetic clusters, five of which correspond to major geographic regions, and subclusters that often correspond to individual populations."	Racial and/or Ethnic Group Differences are "Real" (i.e. Genetic) [31]
Mental Illness	Gardner, M., A., et al. (2006). "Extreme population differences across Neuregulin 1 gene, with implications for association studies." <i>Molecular Psychiatry</i> 11(1): 66-75.	"... allele differences are especially relevant in two SNPs located in a large intron of the gene, as shown by the extreme F_{ST} values, which reveal genetic stratification correlated to broad continental areas."	Populations Differ Significantly in Schizophrenia Susceptibility [32]
Natural Selection	Mekel-Bobrov, N., et al. (2005). "Ongoing adaptive evolution of ASPM, a brain size determinant in <i>Homo sapiens</i> ." <i>Science</i> 309(5741): 1720-1722.	".. one genetic variant of ASPM in humans arose merely about 5800 years ago and has since swept to high frequency under strong positive selection. These findings... suggest that the human brain is still undergoing rapid adaptive evolution."	Brain Size has Evolved More Rapidly in Non-African Populations [33]

Fullerton and Lee *BMC Medical Ethics* 2011 **12**:16 doi:10.1186/1472-6939-12-16

OPEN DATA

Potential Harms to Data Producers' Interests

- First Publication Interests
- Property/Ownership Claims over Methods
 - More of a concern in environmental health research than in other areas
 - What works in genomics may not work elsewhere

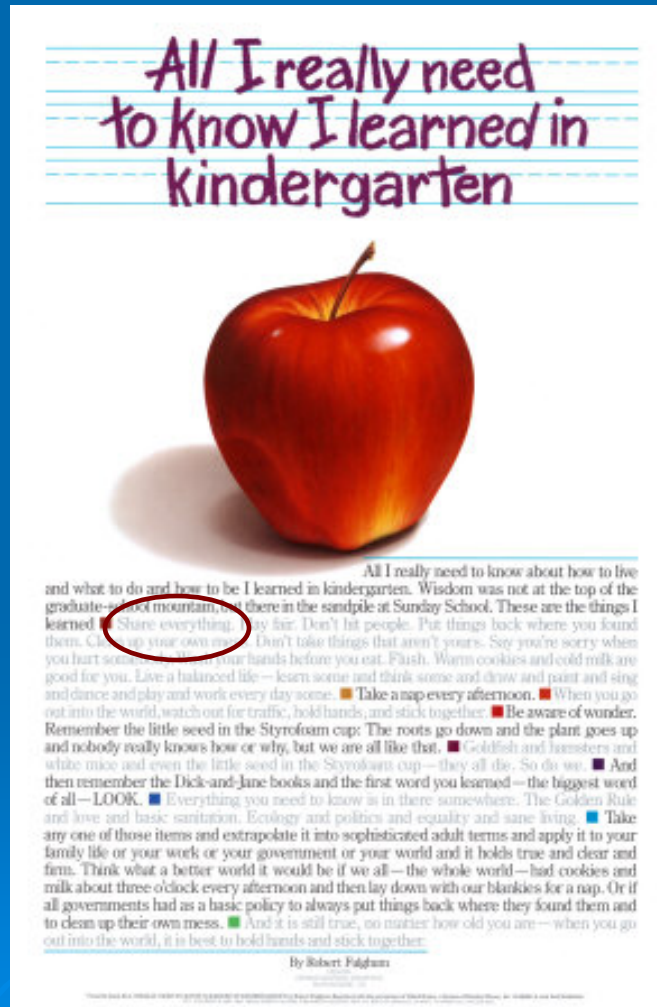


Approaches that Work

- Data Access Committees
- Informed consent: Shift from promising anonymity to consenting to public access, restricted access, or no access
 - People who know more about public data sharing are less likely to consent to research (McGuire 2012)
- Research is needed to determine how researchers interpret and adopt data sharing and withholding policies

Guiding Principles

“Share Everything”



Guiding Principles

- Transparency to research participants must be a priority.
- Enforcement of compliance data-sharing policies should protect original researchers' interests.
 - Researchers need to trust that system works
- DAC and IRB relationships need to be clarified.
- Governance structures should be established to maintain some level of public involvement in decisions about data sharing (not feasible to continue engagement with each participant over extended periods of time).
- Data sharing cannot be an unfunded mandate.

Acknowledgements

This work was supported in part by grants from the National Human Genome Research Institute (P50 HG003390 and R01 HG006281).

We thank Sandra Applebaum, Jalayne Arias, Genevieve Pham-Kanter, Roz Pierson, Karen Seashore Louis, Darren Zinner and Eric Campbell for their help.



Ripple Effects of Data Sharing: Ethical Considerations

Richard R. Sharp, PhD

Director of Bioethics Research, Cleveland Clinic

Co-Director, Center for Genetic Research Ethics, Case Western Reserve

Jalayne J. Arias, JD

Fellow, Cleveland Fellowship in Advanced Bioethics, Cleveland Clinic